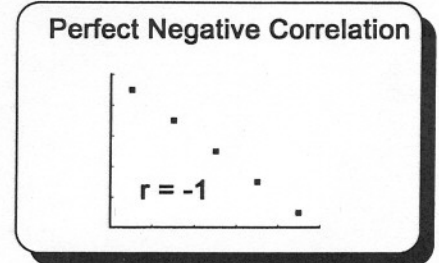# Quick Questions 23 Correlation Analysis
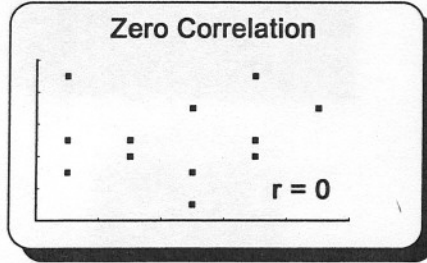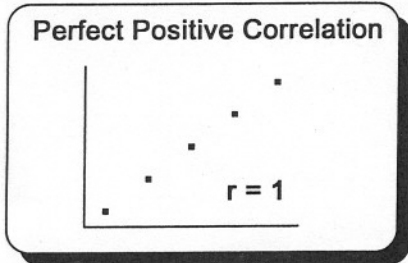
I. Place the number of the appropriate formula, expression, or term next to the appropriate concept.
   A. Coefficient of determination __4__
   B. Coefficient of correlation __2__
   C. A range for r __5__
   D. Coefficient of nondetermination __1__
   E. The test statistic (t) used to measure the significance of r __3__

1.	$1 - r^2$, the variability in y that is not explained by x
2.	$\dfrac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n(\sum X^2) - (\sum X)^2][n(\sum Y^2) - (\sum Y)^2]}}$
3.	$\dfrac{r - \rho}{\sqrt{\dfrac{1 - r^2}{n - 2}}}$
4.	r^2, the variability in y that is explained by x
5.	$-1 \le r \le +1$

II. Draw the following scatters and place the appropriate value for r in the space provided.

Perfect Positive Correlation

r = 1

Zero Correlation

r = 0

Perfect Negative Correlation

r = -1

III. Draw a scatter diagram showing how hours studying per weekend affect grade point average.

Scatter Diagram of Hours Studying and Grade Point Average

$\hat{y}_{\cdot x} = 1.81 + .282x$
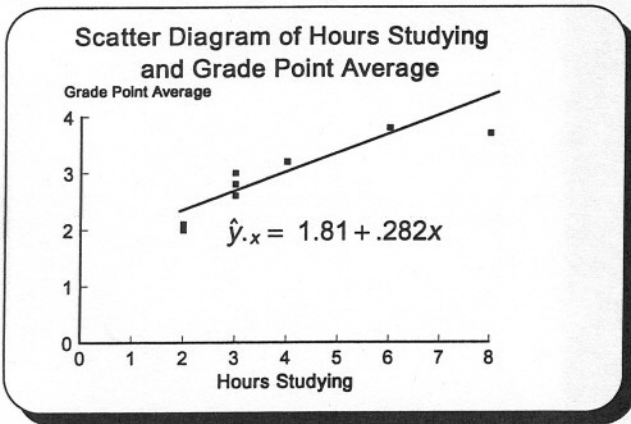
Hours Studying per Weekend(x)	Grade Point Average (y)	XY	X²	Y²
3	3.0	9.0	9	9.00
2	2.0	4.0	4	4.00
6	3.8	22.8	36	14.44
3	2.6	7.8	9	6.76
4	3.2	12.8	16	10.24
8	3.7	29.6	64	13.69
2	2.1	4.2	4	4.41
3	2.8	8.4	9	7.84
31	23.2	98.6	151	70.38

IV. Using the data in question III, calculate the following:
   A. Coefficient of correlation (to 3 decimal places)

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n(\sum X^2) - (\sum X)^2][n(\sum Y^2) - (\sum Y)^2]}} = \frac{8(98.6) - (31)(23.2)}{\sqrt{[8(151) - (31)^2][8(70.38) - (23.2)^2]}} = \frac{69.6}{\sqrt{(247)(24.8)}} = .889$$

   B. Coefficient of determination $r^2 = (r)^2 = (.889)^2 = .790$ or 79.0%

   C. Coefficient of nondetermination $\bar{r}^2 = 1 - r^2 = 1 - .790 = .210$ or 21.0%

   D. Interpret the answer to question IV B.   Seventy-nine percent of grade variability is accounted for by study hour variability.

V. Could ρ (rho) be zero at the .01 level of significance?
   1. The null hypothesis and alternate hypothesis are $H_o : \rho = 0$ and $H_1 : \rho \ne 0$.
   2. The level of significance will be .01 for this two-tail problem with n - 2 degrees of freedom.
   3. The relevant statistic will be r.
   4. If t from the test statistic is beyond the critical value of t, the null hypothesis will be rejected.
   5. Apply the decision rule.

   $df = n - 2 = 8 - 2 = 6 \rightarrow t = 3.707$

   $$t = \frac{r - \rho}{\sqrt{\dfrac{1 - r^2}{n - 2}}} = \frac{.889 - 0}{\sqrt{\dfrac{1 - (.889)^2}{8 - 2}}} = 4.76$$ Reject H_0 because 4.76 > 3.707. Rho could not be zero.

QQ 150 and 151